Ícone Editora Ltda.
R. Rua Anhanguera, 66
Fones: (011) 826-8849 / 826-7074
01135 - S.Paulo

Obra em co-edição com a

EDITORA DA UNIVERSIDADE ESTADUAL DE CAMPINAS (UNICAMP)
Reitor: Paulo Renato Costa Souza



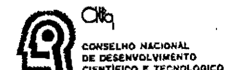
Rua Cecilio Feltron, 253
Cidade Universitária – Barão Geraldo
Fone: (0192) 39.1301 (ramais 2568 e 2585)
13083 – Campinas – SP.

Simon Schwartzman
Claudio de Moura Castro
(organizadores)

PESQUISA UNIVERSITÁRIA EM QUESTÃO







Capa: J.L. de Paula Jr.

Revisão editorial:
Patrícia Campos de Souza e Ilana Pisnky

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA CENTRAL – UNICAMP

Schwartzman, Simon
Sch95p Pesquisa universitária em questão / Simon Schwartzman e Cláudio de Moura Castro. (okg)
Campinas : Editora da UNICAMP, Ícone Editora, São Paulo - CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, 1986.
1. Universidade – Educação superior – Pesquisa. I. Castro, Cláudio de Moura. II. Título.

19. CDD-378.100 72

Índice para catálogo sistemático:
1. Universidade : Educação superior : Pesquisa 378.100 72

ÍCONE EDITORA LTDA.
Rua Anhanguera, 56/66
Fone: (011) 66.3095/826.9510
01135 – São Paulo – SP – Brasil

SUMÁRIO

de um trabalho oneroso em termos de equipes técnico-administrativas, requerendo pessoas de excelente nível e sensibilidade. Por outro lado, uma agência não pode expandir muito essas equipes, sob pena de criar um processo pesado e ineficaz. O tratamento desses grupos deve ser mais a exceção do que a regra.

Em suma, diante de grupos mais maduros e já bem estabelecidos, é necessário também criar condições especiais que ofereçam os recursos e o apoio de que necessitam os mais novos, que contêm a promessa de desenvolvimento. A concorrência pura e simples dificilmente permitiria o progresso desses grupos iniciantes. Mas deve-se entender que o número de grupos com esse potencial é apenas uma fração do total de grupos ou programas em operação. Nem existem disponíveis os recursos materiais ou financeiros para apoiar a todos, nem haveria, em grande parte, o potencial para se beneficiar de tal apoio. Mais ainda, as formas institucionais de apoio tendem a ser artesanais e muito baseadas em tratamento de exceção; não se prestam para expansão indefinida.

Notas

[1] Exceções seriam Brasília e os "filhotes" da USP, as Universidades de Campinas e São Carlos.

II. A QUESTÃO DA QUALIDADE

Claudio Moura Castro

1. Os mistérios da qualidade

"Qualidade... sabemos o que é, todavia, não sabemos. Eis uma contradição em termos. Mas algumas coisas *são* melhores que outras, isto é, têm mais qualidade. Mas quando tentamos dizer o que seria esta qualidade, separada das coisas que a contém, desmorona-se tudo! Não há sobre o que falar. Mas se não podemos dizer o que é qualidade, como saberíamos o que é ou como saberíamos se sequer existe? Se ninguém sabe o que é, então, para todos os efeitos não existiria de todo. Mas para todos os efeitos *existe.* Em que mais se baseiam as notas escolares? Por que outras razões algumas coisas custam uma fortuna e outras vão para o lixo? Obviamente, algumas coisas são melhores que outras... mas o que é 'ser melhor'? E assim seguem rodando as engrenagens mentais sem que encontremos alguma coisa que lhes dê tração. Que diabo é qualidade?" (Pirsig, 1975).

Esta citação de Pirsig freqüentemente adorna, como epígrafe, as discussões sobre avaliação da ciência. Aí estão, a descoberto, as perplexidades do termo. Parafraseando Santo Agostinho, ao comentar o conceito de tempo, não há nada que se possa dizer para quem não tem idéia do que seja qualidade; mas, para quem tem alguma, há muito para se conversar.

O termo é ameaçador quando isolado e colocado na berlinda. Contudo, diante de casos concretos, temos sempre uma idéia do que contém mais ou menos "qualidade". No fundo, tendemos a não ser tão relativistas quanto às vezes professamos, ao afirmar a existência da "nossa" e da "vossa" qualidade. Achamos, enfim, que a "nossa" qualidade é melhor, isto é, a "nossa" qualidade tem mais qualidade!

A qualidade da ciência é o que os cientistas "de qualidade" acham

que é qualidade. Fica por ser melhor trabalhada a epistemologia da qualidade. E ficamos nós com a modestíssima noção, de que as discussões usuais sobre a qualidade das pesquisas tomam como caixa preta epistemológica as razões e fontes últimas desse julgamento. Busca-se tão somente mapear consenso e dissenso nas percepções daqueles que foram longamente preparados para a ciência e cuja experiência cotidiana gera intimidade com os assuntos sob julgamento.

E é com tão magra dieta epistemológica que prosseguimos. Não tanto por gosto como pela escassez de alternativas.

2. A avaliação da ciência: os métodos usuais

Quanto mais ativo e produtivo o ambiente científico, mais freqüentes e rigorosas são as rotinas de avaliação. Nos países que lideram o mundo da ciência, cultiva-se prodigamente um emaranhado sistema de apreciação de propostas, instituições, grupos, pesquisas, cursos etc. Em outras palavras, parte considerável dos financiamentos transitam por mecanismos onde a avaliação tem grande peso.

Poderíamos mesmo perguntar se os mandamentos da dúvida sistemática, da discussão aberta e da crítica frontal, considerados ingredientes essenciais do processo científico, não teriam estimulado esta propensão da ciência organizada para os infindáveis ritos de avaliação. De fato, é difícil pensar em outra área onde tanto esforço seja dedicado à avaliação e onde suas conseqüências sejam levadas tão a sério.

Todavia, o vínculo entre avaliação e financiamento, hoje tão claramente registrado, esconde uma história deveras curta. É somente a partir da década de 60, quando os recursos para a ciência e tecnologia tornam-se mais escassos, que os mecanismos de avaliação passam a florescer. O que tomamos como tradicional no financiamento à pesquisa não passa, portanto, de procedimentos há pouco adotados.

Mais ainda, não podemos supor que essas práticas sejam iguais entre países, disciplinas ou instituições, ou que sejam estáticas. E já que nossa incipiente convivência com a avaliação acaba, de uma forma ou de outra, sendo influenciada pela experiência dos países cientificamente mais expressivos, parece apropriado explicitar como esses países procedem para avaliar suas pesquisas.[1]

O início da avaliação como prática um pouco mais sistemática se dá no pós-guerra com o financiamento à pesquisa e a contratação de projetos pelas agências públicas. Em última análise, sua origem é militar. Mais especificamente, é o Office for Naval Research norte-americano que formaliza esses procedimentos (Roy, 1983, p. 40).

O processo decisório, nestes casos, cabia ao pessoal técnico e administrativo da própria instituição. Progressivamente, no entanto, passou a haver uma participação maior de pares da comunidade científica. Não obstante, ainda é comum e respeitável, em todos os países, a utilização desses formatos onde predominam as equipes intramuros da agência patrocinadora.

Há duas grandes vertentes nos processos de avaliação. A mais usual é a avaliação pelos pares, fortemente ancorada na reputação adquirida pelo avaliado. A segunda deriva de critérios mais quantitativos, desembocando na bibliometria e cientometria. Previsivelmente, as alternativas mais interessantes parecem residir em combinações dos dois métodos.

3. A avaliação por pares

A avaliação pelos pares (*peer review*) baseia-se no conhecimento acumulado pelos cientistas considerados em condições privilegiadas para julgar as pesquisas e instituições de sua área. A experiência demonstra que esse é um método simples, barato, suficientemente seguro e aplicável a uma gama bastante ampla de situações.

Em geral, cientistas tendem a julgar conscienciosamente seus colegas. Tomadas as precauções usuais de aumentar o número de juízes e de promover, com freqüência, a sua renovação, a margem de erro é, na maioria dos casos, suficientemente reduzida. Precauções adicionais são normalmente adotadas. Muitas vezes é mantido o sigilo dos juízes para que sobre eles não pesem acusações de favoritismo ou discriminação e, sobretudo, para que não haja pressões diretas, insinuando ou instando certos resultados. Mas, na verdade, o sigilo é uma faca de dois gumes, pois desonera o juiz da responsabilidade tanto pelos acertos politicamente difíceis quanto pelos erros e tendenciosidades.

Há várias maneiras de lidar com esse problema. O aumento do número de juízes, embora permita reter um sigilo externo, torna mais visível para os outros pares a presença de elementos de julgamento estranhos ao assunto. Certamente não se pode esperar confidência diante de situações em que se suspeita de falta de lisura.

Outra alternativa, é o chamado procedimento *double-blind*, no qual não só os juízes permanecem incógnitos, mas também os trabalhos perdem a sua identificação. Há excelentes razões para a sua adoção em grandes programas e concursos, em que se lida com amplos números. Porém, sua aplicação em áreas pequenas é inviável, pois o mero assunto, a escolha da bibliografia e as técnicas utilizadas praticamente identificam o autor, tanto quanto o teor do laudo aponta o juiz.

Uma restrição adicional ao método de avaliação por pares é o seu caráter de refletir um conhecimento compartilhado na comunidade científica, onde, como em qualquer outro grupo, circulam mitos e enganos. O mais comum deles é o chamado efeito halo: pode continuar sendo conside-

rado bom o que já deixou de sê-lo e é considerado bom o que está próximo do que comprovadamente o é. As reputações persistem apesar de já ter-se deteriorado o desempenho. As "vacas sagradas" são muitas vezes julgadas mais pelo que fizeram do que por sua real probabilidade de realizarem o trabalho em pauta. E o halo da excelência transborda para as suas adjacências. Lewis Solmon, um experimentado organizador de avaliações de cursos de pós-graduação, conta que o Departamento de Administração de Empresas da Universidade de Princeton foi considerado um dos melhores, apesar de Princeton *não ter* cursos de administração.

Utilizando a analogia de Irving e Martin, o juízo dos pares funciona bem nos regimes de "concorrência perfeita" e é falho nos casos de "oligopólio". De fato, o julgamento dos pares funciona corretamente quando há muitos juízes, muitos julgados e quando o objeto do julgamento não é excessivamente dispendioso, de forma a pesar no todo. Mas ocorre que, sobretudo no pós-guerra, a ciência passa a ser realizada em grupos cada vez maiores. A exigência de equipamentos muito caros requer grandes times de pesquisadores.

O crescimento da *big-science* atinge inclusive o Brasil. Aceleradores de partículas, espectógrafos, cromatógrafos, microscópios eletrônicos, grandes telescópios e muitos outros equipamentos absorvem fatias grandes e indivisíveis de recursos. Nessas áreas, praticamente todos os pesquisadores fazem *big-science* – não se trata de uma alternativa de trabalho, mas da única alternativa.

Se assim é, os juízes escolhidos também participam de um determinado programa. E como os recursos que julgam são fundos substanciais no total dos orçamentos, esses juízes têm a ganhar ou perder com o resultado das avaliações. Se o laboratório concorrente ganha recursos, sobrará menos para o seu próprio. Trata-se, portanto, de um julgamento que exige mais do juiz do que seria apropriado esperar. Daí a tendência a se dar maior relevo ao papel dos funcionários governamentais que, pelo menos, não têm comprometimento direto. Mas isso não deixa de trazer outros problemas, que examinaremos mais adiante.

Nesse contexto, cabe enfatizar outro problema que, embora ocorra em todos os níveis, é mais grave na *big-science*. Aprovar um recurso significa dar ao grupo agraciado recursos adicionais, mas a decisão oposta, de interromper o fluxo de recursos, pode ter conseqüências nefastas para o recipiente. São decisões que precisam ser tomadas, sobretudo em épocas de crise, quando os recursos se reduzem e os grupos apresentam fraco desempenho. O ônus do juiz que toma essa decisão é particularmente pesado, principalmente nas áreas pequenas e nas de *big-science*, em que se manipulam recursos de grande monta.[2]

É ainda pertinente registrar as dificuldades encontradas para se cortar bolsas de pesquisadores que apresentam fraco desempenho ao longo do

tempo. Essa relutância tem sido observada tanto no CNPq quanto em seu congênere francês, o Centre National de Recherche Scientifique – CRNS, com conseqüências deveras negativas, já que leva a reter pesquisadores improdutivos e subtrai um potente estímulo para o sistema como um todo.

4. As avaliações quantitativas

Diante dessas dificuldades, têm sido propostos critérios mais desvinculados de julgamentos subjetivos e imediatos. No passado recente desenvolveram-se técnicas de apreciação da ciência por métodos quantitativos. Trata-se, na verdade, de manipular quantitativamente julgamentos qualitativos. No fundo, é um processo de contagem de coisas que já foram qualitativamente julgadas. Por exemplo, uma instituição que tem cinco pesquisadores excelentes é considerada superior, como grupo de pesquisa, quando comparada a uma que só tem dois pesquisadores de mesmo gabarito. Dizemos que uma instituição que produz três contribuições importantes ao conhecimento coloca-se melhor do que outra que só produziu uma do mesmo nível.

A biliometria, ou cientometria, nada mais é do que uma técnica para tratar comparações desse tipo de forma mais completa e rigorosa. Seus ingredientes básicos são o reaproveitamento de avaliações geradas para outros fins e as técnicas quantitativas para o seu tratamento.

As vantagens da bibliometria residem na diluição dos elementos de julgamento. Os escores quantitativos constituem o somatório de muitos pequenos julgamentos e apreciações, realizadas por dezenas de pessoas.

Todo produto da ciência acaba sendo comunicado de forma escrita. Assim, ao se monitorar a ciência no seu momento de publicação, temos uma imagem bastante adequada do seu fluxo. Essa afirmativa é correta se entendemos ciência numa acepção muito estrita, isto é, como algo que se distingue da tecnologia, pois, em se tratando de tecnologia, a publicação pode mesmo ser indesejável. É de fato impossível discutir a avaliação da ciência e da tecnologia conjuntamente.

Na prática, o Brasil desenvolveu uma política de apoio à ciência muito mais enérgica e eficaz do que conseguiu fazer em relação à tecnologia. Talvez isto tenha sido mesmo inevitável. É mais fácil produzir ciência do que tecnologia. Ciência se produz para outros cientistas, quase que em círculo fechado. Tecnologia só faz sentido à medida que hajam, no campo econômico, decisões de fazer uso do produto. É algo muito mais difícil e que talvez só adquira real viabilidade uma vez feitas certas conquistas preliminares na institucionalização da ciência.

Muito do que se aprende no desenvolvimento de uma comunidade científica pode ser reaproveitado na tecnologia. Mas nem tudo. Tratam-se de produtos diferentes. Na ciência o produto é o trabalho escrito, cuja qua-

lidade é inicialmente aferida pelo prestígio de quem a publica. Na tecnologia a publicação é o acessório: o importante é o produto.

A aplicação à tecnologia dos sistemas de estímulos usados na ciência pode levar à publicação de coisas que ou deveriam ser mantidas sigilosas ou então patenteadas. Pode levar também a um falso direcionamento dos programas, à medida que, por analogia com os procedimentos científicos, esses programas são avaliados pelo volume de suas publicações.

Outro problema é o ritmo de produção. Na ciência, os prazos funcionam como mecanismos de pressão para manter a intensidade de dedicação e uma certa eficiência no processo. Conseqüentemente, sua elasticidade é usual ou mesmo crônica. Já na tecnologia, os prazos são fatais. Os produtores têm programações de implantação de linhas e, como pano de fundo, a presença de seus concorrentes. Se os resultados se atrasam, podem mesmo não interessar mais, além de azedar as relações escola-empresa.

Voltando às publicações, seu acompanhamento pode se dar através de níveis diferentes de sofisticação e seletividade. Trata-se de um processo no qual se criam critérios segundo os mais variados graus de exigência. Por exemplo, pode-se incluir tudo que foi impresso, pode-se excluir revistas de circulação interna, pode-se exigir corpo eleitoral rígido para os periódicos escolhidos e assim por diante.

Nos países cientificamente maduros, o aperfeiçoamento dos indicadores leva a critérios mais rígidos de inclusão – são arrolados apenas os artigos publicados nos periódicos com critérios mais competitivos. Na prática, são arrolados os artigos contidos em listas restritivas de periódicos, como o *Current Contents*, que inclui cerca de 3.000 periódicos, todos considerados exigentes nos critérios de publicação e dotados de corpos editoriais formados por cientistas renomados e oriundos de instituições diferentes.

O passo seguinte é tomar não a publicação diretamente, mas o número de vezes em que é citada. A fonte usual, nesse caso, é o *Citation Index*. Esse indicador introduz um controle de qualidade adicional, pois requer que a publicação tenha sido lembrada por alguém que assina artigos em periódicos de qualidade, listados no sistema de referência (na prática, o mesmo *Current Contents*).

Finalmente, visando captar não a quantidade de publicações razoavelmente boas, mas aquele número restrito de contribuições excepcionais que determinam os saltos qualitativos da ciência, costuma-se tomar também aqueles artigos que foram citados, digamos, mais de dez vezes.

Com estes três indicadores compõe-se a rede de monitoração da produtividade científica. São computados não apenas os artigos submetidos ao controle de qualidade dos corpos editoriais das revistas de melhor nível, mas também suas repercussões na literatura científica. De um lado, capta-se o volume de produção; de outro, identificam-se as contribuições que receberam aplauso implícito da comunidade científica.

Como se poderia esperar, esse sistema possui falhas e equívocos. Em geral, criticam-se as contagens de publicações por darem igual peso a contribuições de valor diferente. A crítica é correta em princípio, daí a utilização das citações dos artigos como controle adicional. Todavia, deve-se entender que quando comparamos grupos de pesquisadores pela "lei dos grandes números", os erros tendem a se cancelar. Por outro lado, os estudiosos de bibliometria tendem a concordar que qualidade e quantidade estão estatisticamente associadas: as melhores pesquisas originam-se dos locais onde se publica muito (Grasberg, 1959, pp. 31-37; Carakushamsky, 1982, p. 2 e Narin, 1976, p. 2).

Outra crítica usual e perfeitamente legítima é a de que cada área de ciência tem padrões diferentes de publicação. Matemáticos publicam pouquíssimo, biólogos prodigamente e assim por diante. Desse modo, são precaríssimas as conclusões que se podem derivar de comparações entre diferentes ramos das ciências a partir de contagens de publicações.

A análise das citações tem também os seus problemas. No *Science Citation Index*, apenas o primeiro autor é citado; ademais, há problemas de homônimos e erros de identificação. Existem também fenômenos espúrios: certos artigos são muito citados por conterem algum erro sério, por estarem associados a algum método ou serem alvo de severas críticas. Mais comuns são os efeitos de halo, que levam à citação, por prestígio, de autores que já são mais conhecidos ou que trabalham em centros grandes, portanto, próximos de muitos colegas que podem citá-los. Esta é a chamada Lei de Mateus: A quem já tem, mais lhe será dado.

Esses não são problemas inventados ou artificiais. Eles existem e devem ser levados a sério, dentre outras razões, pela possibilidade de se reduzir a eliminar alguns.

Todavia, os autores que mais vêm lidando com esses assuntos não crêem tratar-se de limitações que comprometam a utilidade desses indicadores como instrumentos de avaliação de grupos de pesquisa.

Cumpre, nesse momento. trazer a discussão mais próximo ao nosso meio. De que nos serve isso tudo?

Há inicialmente que comentar as fontes de informações que nos poderiam servir. Como é sabido, o *Current Contents* inclui poucos periódicos brasileiros. Já o *Citation Index* inclui todas as publicações citadas nas revistas listadas. Na prática, menos de 10% das revistas científicas brasileiras identificadas por levantamentos da CAPES e do IBICT aparecem no *Citation Index*, o que significa que, se este índice capta bem o que circula fora do Brasil ou entre os autores das poucas revistas brasileiras indexadas, ele deixa de lado a massa de trabalhos que só circulam em âmbito nacional (ver Capítulo III da Terceira Parte deste volume.)

A melhor informação sobre nossa produção científica universitária vem dos relatórios da pós-graduação encomendados pela CAPES. Trata-se de um dado bruto, incluindo todos os títulos citados pelos cursos. Os con-

troles de qualidade são estabelecidos *a posteriori*, através de critérios de avaliação dos periódicos citados. Todavia não há, em um horizonte de tempo razoável, qualquer possibilidade concreta de gerar um índice de citações.

Se a inexistência de tal índice para a literatura brasileira não incluída no *Current Contents* é uma limitação séria, um outro aspecto torna a questão algo distinta. Na verdade, grande parte dos objetivos atendidos por contagens deste tipo são estritamente comparativos. Contar publicações nos permite dizer quem produz mais artigos que atingem certos limiares de qualidade. Nos locais de liderança mundial da ciência, onde o fluxo de publicações é abundante, o poder de discriminação dos indicadores aumenta com o uso de critérios mais rígidos de inclusão. Nesses casos, observamos que a contagem do total de publicações não se correlaciona com outros indicadores de excelência.

Mas no caso do Brasil, onde é muito mais ralo o fluxo de publicações, temos que pensar em indicadores com poder discriminatório em diferentes patamares de desempenho. É apropriado, por exemplo, compararmos a Economia da PUC/Rio e a da USP pelo número de artigos em revistas citadas no *Current Contents*. Esse talvez seja o indicador mais discriminante para esses cursos, que estão dentre os melhores do país. Não obstante, examinando os cursos mais modestos e, principalmente, os menos consolidados, notamos que sua contribuição para estes mesmos periódicos é nula, com raras exceções. Já o número total de publicações – pesadamente influenciado pelas publicações em periódicos locais e sem o menor controle de qualidade – é significativamente diferente de zero em qualquer um deles. É provável que, nesse patamar, seja melhor centro de pesquisa aquele que produz significativamente mais publicações de qualquer qualidade. A utilização de um indicador mais seletivo excluiria justamente as publicações que discriminam esses cursos.

É de se notar, contudo, o uso reticente que é dado a esses métodos quantitativos. Voltaremos a discuti-los, mais adiante, quando examinarmos o sistema de avaliação da CAPES, a única instituição a tratá-los de forma sistemática.

Ao concluir esta seção, vale lembrar que a avaliação por pares e a cientometria são casos polares. Na verdade, os exemplos mais interessantes de avaliação combinam, de alguma forma, os dois métodos.

5. A experiência internacional

Antes de examinar alguns casos ilustrativos, vale a pena discutir questões relativas ao grau de institucionalização do processo avaliativo. De fato,

há processos de avaliação que, de tão informais e mesclados com outros procedimentos, podem passar despercebidos. Por exemplo, fundações filantrópicas como a Fundação Ford e a Fundação Rockefeller utilizam como administradores das suas dotações de pesquisa, indivíduos cujo currículo acadêmico nada os inferioriza em relação aos pesquisadores financiados. Para todos os efeitos, são pares aceitos como tais, podendo mesmo, em conversações informais, oferecer sugestões técnicas de valia. Ao mesmo tempo, essas entrevistas permitem formar juízos corretos e explícitos do andamento da pesquisa.

No outro extremo, temos os procedimentos utilizados na seleção dos artigos a serem publicados nos periódicos mais respeitados. Aqui, julgador e julgados não conhecem suas respectivas identidades. O parecer é enviado pelo editor da revista, obliterando a assinatura do autor. O parecerista, por sua vez, recebe um original não assinado. Procedimento equivalente é utilizado nos concursos de ensaios e monografias.

Em boa medida, o grau de institucionalização dos processos de avaliação tem a ver com o que se avalia. Avaliar o processo científico é algo vago. Na prática, avaliam-se pesquisadores, projetos de pesquisa, grupos, departamentos, cursos, propostas etc. Ademais, pode-se avaliar a proposta ou o resultado, isto é, *ex ante* ou *ex posto facto*.

Mesmo nos países mais avançados cientificamente, a maioria dos projetos de pesquisa são avaliados *ex ante*. Isso tem sido objeto de fundadas críticas, já que o produto final é mais fácil de ser avaliado e o laudo obtido mais confiável..

A tendência mais usual é dar um grau maior de formalização às avaliações individuais, seja de pesquisadores, seja de propostas de pesquisa. Este é o sistema dos comitês de consultores que recebem propostas dentro de normas padronizadas. Já os grupos de pesquisa e as instituições podem, em muitos casos, ser avaliados sem periodicidade fixa, sem formatos pré-determinados e por juízes recrutados *ad hoc*. Há igualmente diferenças associadas a países. A Inglaterra tem uma tradição mais longa de avaliação, mas adota geralmente procedimentos pouco formalizados. Já nos Estados Unidos há uma superabundância de sistemas institucionalizados, mas nem sempre públicos, com comitês de pares e farto uso de bibliometria.

Examinamos em seguida alguns exemplos de avaliações, escolhidos pela sua importância e variedade. Nosso interesse é não só apresentar exemplos que ilustrem a gama de modalidades de avaliação, como também ilustrar diferenças no objeto da avaliação, já que são julgados tanto projetos como instituições. Mostramos inicialmente o sistema americano de avaliação de cursos de pós-graduação. Depois, os mecanismos de seleção de projetos de pesquisa individuais na National Science Foundation, seguido de um sistema holandês, também de seleção de projetos. Finalmente, discutimos um interessante método comparativo para a avaliação de programas ou instituições individuais.

6. As avaliações da pós-graduação americana

A primeira tentativa de avaliação conjunta das universidades americanas deve-se a Hughes e data de 1928. Todavia, os grandes avanços foram feitos por Alan Cartter em 1966 e Roose & Anderson em 1970. Desde então, inicia-se uma seqüência pródiga e pluralista de avaliações de cursos e universidades, tomando conhecidos nomes como Astin, L. Solmon e muitos outros, associados a uma tradição de avaliações baseadas na reputação dos cursos. O artigo de Judith Lawrence e K. Greene (1980, p. 17) contém uma bibliografia com 221 títulos de artigos e livros sobre a experiência americana com análises de reputação – alguns apresentando avaliações, outros criticando ou comentando (Lawrence, 1980).

A essência desses procedimentos consiste em pedir a um grupo de professores universitários que coloquem cursos ou universidades em ordem decrescente de sua excelência, tal como cada um as percebe. Estas classificações são consolidadas em escalas que permitem hierarquizar os cursos.

Note-se que há dois traços marcantes nesse processo. Em primeiro lugar, os professores não têm acesso a qualquer material ou informação adicional sobre o que estão julgando. Sua apreciação refletirá estritamente a reputação que formaram previamente. O método visa tão somente captar uma idéia consolidada de hierarquia de reputações.[3]

Em segundo lugar, os painéis de avaliação são formados por professores típicos da média e não por aqueles que poderiam ter sido escolhidos pelo maior peso de sua opinião. O interesse não está no julgamento do cientista mais destacado, mas no do cientista médio.

É de se notar o elevado grau de consistência encontrado nas comparações entre hierarquias geradas em pesquisas diferentes, bem como com outros indicadores de excelência acadêmica, o que sugere que as medidas captam dimensões estáveis e bastante invariantes em relação ao que medem. Apesar de todos os aperfeiçoamentos e modificações introduzidas ao longo do tempo, o resultado surpreendente é que as instituições situadas no topo da pirâmide pouco mudam. Praticamente as mesmas 50 instituições reaparecem com 58 anos de intervalo.

As medidas que buscam coisas como a excelência acadêmica, o prestígio e a produtividade científica dos programas podem ter correlação entre si de até 0,99. Em outras palavras, há um conjunto de dimensões de qualidade e *status* que parecem ser altamente convergentes (Lawrence e Green, 1980, pp. 3, 19 e 20). Esse é o caso, sobretudo, do volume de publicações, do prestígio dos professores e dos indicadores de recursos materiais e humanos.

No decorrer das muitas avaliações realizadas foi possível identificar importantes fontes de distorções e ruídos, até certo ponto evitáveis: ex-alunos supervalorizam suas universidades de origem; professores supervalorizam suas instituições; departamentos muito grandes e antigos têm mais

probabilidade de ter seus ex-alunos nos painéis de avaliadores; instituições mais distantes dos tradicionais centros universitários são menos conhecidas e, portanto, subestimadas.

O próprio Cartter, o mais cuidadoso e experimentado desse grupo, advertia acerca do caráter disciplinar desse processo de avaliação. Só é possível, na realidade, avaliar-se cursos ou departamentos. Assim, as "químicas" estatísticas, que permitem agregar dados para avaliar universidades como um todo, não mereceriam qualquer confiança. Mas nem por isso essas agregações deixam de ser feitas.

O mais grave é que somente os melhores programas aparecem. Outros menos destacados sequer são julgados ou seus resultados não merecem confiança. Os resultados, portanto, servem mais para referendar as posições tradicionais de um clube privilegiado do que à maioria. Nesse sentido, têm um efeito conservador de preservar o *status quo*, já que esses resultados terminam tendo peso na alocação de recursos e escolhas dos estudantes.

É preciso também não se iludir com o aparente rigor dos números. Entre o primeiro, o segundo e o terceiro cursos, pouca diferença real se pode atribuir. Apenas pode-se afiançar verdadeiras diferenças quando há uma significativa distância entre duas instituições.

Outro ponto de magna importância é saber o que se está medindo. Em geral, tais estudos medem a excelência da pesquisa produzida e a reputação dos professores. Sobre a qualidade do ensino nada se diz, exceto em alguns poucos estudos em que isso é explicitamente buscado.

Além de nada dizer sobre instituições menores, toda tradição e esforço avaliativo desses estudos se concentram na pós-graduação. Na verdade, portanto, é mais uma avaliação da pesquisa que do ensino.

À guisa de conclusão, podemos dizer que desse enorme esforço americano deriva um conjunto de resultados em que se sobressai a forte consistência na identificação dos mais destacados grupos de pesquisa. São resultados importantes e muito úteis, desde que neles não se busque mais do que isso. Com eles ficamos sabendo quem chegou ao topo e nisso tem-se confiança. Em particular, avaliam-se apenas os melhores, apenas a pesquisa e apenas a pós-graduação.

7. O modelo da National Science Foundation

Sendo a principal fonte de financiamento para pesquisas do país de maior produção científica do mundo, a National Science Foundation acaba tornando-se uma referência inevitável. É interessante notar que seu corpo de funcionários permanentes é muito pequeno: alguns funcionários administrativos e um grupo pequeno de renomados cientistas. O julgamento das propostas individuais, nessa instituição, é realizado por juízes externos, cientistas de boa reputação. Esses cientistas recebem propostas na sua

especialidade ou subespecialidade, emitindo um parecer e uma nota final. Os pareceres são reunidos e examinados em seu conjunto pelos "cientistas-burocratas" da própria NSF, que emitem um julgamento final, aprovando ou não a proposta, isto é, confirmando ou não o parecer conclusivo dos juízes. (Em algumas áreas esses procedimentos são complementados por painéis externos de cientistas, que também opinam sobre as propostas). Mas sabe-se que quase sempre o "pesquisador-funcionário" da NSF ratifica o julgamento do consultor externo.

Há um ponto metodologicamente crítico nesse processo. Para que os julgamentos individuais sejam comparáveis, é necessário que os critérios dos consultores externos tenham o mesmo nível de rigor ou exigência. Naturalmente, o papel dos juízes internos da NSF é fazer essa compatibilização. Todavia, isso só funciona se a descalibragem entre consultores for apenas ocasional. Se esta fosse endêmica ou crônica, as notas de nada serviriam. Restaria a análise técnica e a substância do parecer, mas o volume de trabalho para o exame de cada caso tornaria inviável a operação do programa sem grandes equipes.

E de se notar que, de fato, parece haver consistência nos padrões dos juízes, o que pode ser verificado pela elevada taxa de confirmação dos julgamentos dos consultores externos.

Contudo, é também oportuno registrar que justamente o mais central e visível sistema de avaliação por pares tem sido objeto de acusações candentes, gerando uma movimentada controvérsia, centrada nos editoriais da revista *Science.*

R. Roy enfatiza, além dos pontos tratados anteriormente nesse ensaio, as questões de escolha dos pares, a falta de possibilidade de apelo, a elevada margem de erro do sistema e o excesso de tempo mobilizado para preparar propostas de pesquisa. Vale a pena rever esses pontos, por se referirem a questões que em muito extrapolam problemas peculiares à National Science Foundation.

Em pesquisa recente, Cole, Cole e Simon (1981) mostram que há uma margem de erro da ordem de 25% nas decisões da NSF. Por virem de um trabalho metodologicamente bem cuidado, números tão elevados provocam considerável mal-estar, tal como o documentado pela pródiga troca de cartas da *Science.*

Pela análise dessa pesquisa, parece que é na "loteria" da escolha de consultores que residem as divergências de laudos, já que, nos seus procedimentos, Cole e seus associados submetem os projetos julgados pela NSF a novo julgamento, por outros pares, derivando daí a margem de 25% de desencontro. De fato, há o problema de tradução diferente das preferências — "bons" e "excelentes", por exemplo, são duas maneiras legítimas de mostrar aprovação, mas podem, no cálculo final das médias, significar a diferença entre aprovação e rejeição. Todavia, como os próprios autores indicam, se realmente há uma margem de enganos e equívocos resultantes

da compreensão inadequada de instruções, esta não é a fonte principal das divergências. No fundo, há diferenças de percepções acerca do que é boa ciência e acerca de prioridades.

Na verdade, a posição de Cole não pode ser tomada simplesmente como uma acusação ao sistema de avaliação por pares, mas, sobretudo, como uma demonstração de que o consenso acerca do que é prioritário ou do que é boa ciência é inevitavelmente parcial. A ciência trata com o desconhecido e, conseqüentemente, deve conviver com uma área de penumbra nas suas decisões táticas e estratégias.

Por outro lado, é necessário contextualizar esses resultados. Tal margem de erro é também função da natureza dos candidatos às dotações do NSF. Quanto mais variado for o nível de qualificação dos candidatos que se apresentam, menor a margem de divergências entre juízes diferentes. Levando ao extremo — apenas para facilitar a compreensão — entre uma proposta de um Prêmio Nobel e outra de um estudante de uma instituição medíocre, é de se imaginar que os juízes tenham pouca dificuldade de concordar na identificação da melhor. As dificuldades surgem nas avaliações de projetos de nível intermediário.

Podemos pensar que se esses mesmos juízes da NSF estivessem avaliando propostas, digamos, do CNPq, a percentagem de erro seria menor do que os 25%. Não há dúvidas de que os nossos pesquisadores apresentam uma margem maior de propostas pouco convincentes ou com equívocos facilmente detectáveis, isto é, que a dispersão da qualidade é maior. As divergências entre juízes, nesse caso, concentrariam-se na faixa das pesquisas aprovadas, não tendo portanto conseqüências práticas.

Contudo, como aponta R. Roy, não devemos deixar de ver com cautela a dimensão de arbitrariedade presente no processo de escolha dos consultores. Já que "cada cabeça é uma sentença" e selecionar corretamente os consultores colabora para a obtenção dos resultados desejados, esse processo de escolha deve ser objeto de muito cuidado.

Roy traz também à discussão o excesso de tempo requerido dos cientistas para preparar projetos. De fato, se há divergência de opiniões, quem apresenta mais propostas maximiza as chances de ser financiado. Suas estimativas indicam que os melhores cientistas americanos gastam entre um quarto e metade do seu tempo ou preparando ou julgando propostas (mais próximo da metade para os mais eminentes). Esse argumento, na verdade, é menos uma acusação ao julgamento por pares do que ao sistema de financiamento de pesquisas baseado em projetos detalhados.

Finalmente, haveria que mencionar a ausência de mecanismos de recurso às decisões dos painéis. Isso não é uma característica necessária do sistema. Contudo, é muito usual, inclusive na NSF. Essa incapacidade do sistema de detectar e corrigir erros grosseiros é geralmente exemplificada com o caso do pesquisador que teve sua proposta recusada por não incluir entre os signatários um famoso pesquisador do mesmo departamento. A

suspeita de que poderia haver algo errado se este nome não figurasse é, em princípio, justificada. Todavia, a acusação é de que não houve mecanismo de correção para retificar o julgamento negativo, já que tal pesquisador havia morrido dois anos antes!

Não há dúvidas de que tais críticas à NSF não são infundadas. O sistema erra, como devem errar todas as variantes adotadas pelo mundo afora. Acontece, contudo, que outras alternativas não se revelaram melhores. As opções apresentadas na *Science* não escapam de críticas semelhantes.

8. O sistema interativo holandês

A agência holandesa de financiamento à pesquisa lançou um novo e completo sistema de julgamento de propostas que está em operação na área de física pura. Vale a pena analisar suas principais características, já que estas respondem à tentativa de evitar as limitações freqüentemente encontradas nos sistemas mais convencionais.

Além de físicos, o sistema tem como juízes, químicos, matemáticos e astrônomos. Mais ainda, inclui necessariamente dois estrangeiros. A esse grupo de consultores pede-se que comentem as propostas em seus méritos, interesses e limitações intrínsecas. Não lhes é perguntado se o projeto deve ou não ser financiado – só quem vê todas as propostas terá uma visão de conjunto que permita comparações.

Esses pareceres, sem identificação dos autores, são enviados aos proponentes da pesquisa para que tenham oportunidade de comentar ou responder às críticas. De posse de todo esse material, o júri dá notas e hierarquiza as propostas, com base nos objetivos do trabalho, sua metodologia e qualidade.

A decisão final de financiar ou não permanece uma prerrogativa da diretoria da agência. Na realidade, contudo, é raro deixar de haver um consenso entre os membros do júri. Na prática, confirma-se sempre a indicação desse júri.

9. O método dos indicadores parciais convergentes

Irving e Martin, pesquisadores da Science Policy Research Unit (SPRU) da Universidade de Sussex, vêm realizando vários estudos de avaliação de instituições individuais baseados na estratégia de utilizar um número grande de indicadores parciais. Como cada um deles pode estar sujeito a erros, não seria confiável basear julgamentos finais em critérios únicos. Mas, no seu conjunto, eles podem apontar para conclusões que serão tanto mais robustas quanto mais se mostrarem convergentes os diferentes indicadores.

Outra característica desse método é a preocupação de comparar apenas instituições rigorosamente da mesma área e, se possível, que tenham equipamentos comparáveis.

Vale a pena ilustrar seus procedimentos recapitulando um estudo recente desses autores, contratados para avaliar o telescópio Isaac Newton do Observatório de Greenwich (Irving e Martin, 1983).

Sua estratégia inicial foi identificar três outros observatórios que tivessem telescópios óticos de tamanho comparável (Cerro Tololo, Kit Peak e Lick). Em seguida, levantaram todas as publicações de cada observatório no período considerado. Estas listas foram então complementadas com estimativas de produtividade *per capita*, tabulações de citações destas publicações (através do *Citation Index*) e com os artigos citados muitas vezes. Foram também estimados os custos de produção de cada artigo, a partir dos orçamentos das instituições.

Paralelamente, foram ouvidos os pesquisadores da área, que opinaram sobre os méritos e limitações de cada um dos observatórios. Esses pesquisadores foram questionados acerca das diferenças, dificuldades e trajetórias das instituições.

O estudo identificou assim uma significativa desvantagem de Greenwich em relação a vários de seus congêneres. Suas publicações eram muito menos freqüentes e isso não era compensado por um número maior de citações. O custo por publicação também era superior ao dos outros.

Os cientistas consultados confirmaram esse desempenho mais fraco e apontaram suas causas. Revelou-se que, mesmo antes de ser construído, já se sabia ser inadequado o local – praticamente dentro de Londres – pelas más condições de visibilidade decorrentes do clima e da localização. Mais ainda, declararam que, sendo uma "repartição pública", Greenwich oferecia condições de trabalho rígidas e burocratizadas, em comparação com instituições universitárias. Sua instrumentação complementar também era deficiente.

A publicação do relatório gerou uma reação irada do seu antigo diretor, que escreveu à revista *Social Studies of Science* defendendo-se e contra-atacando os dois autores (Smith, 1983).

Não nos interessa, particularmente, a evolução da disputa e sim exemplificar a utilização engenhosa de: 1) comparações entre instituições similares; 2) o uso de publicações e suas citações com o indicador de produtividade; 3) o uso de julgamento de pares. Parece bastante razoável aceitar o aumento de confiança nos resultados quando se observa convergência dos indicadores parciais.

10. A avaliação da pesquisa no Brasil

Para um país de fraquíssimas tradições científicas, o Brasil até que

não está mal em termos de avaliação de pesquisa. Há um bom começo e um aceitável grau de capitalização da experiência que vai sendo adquirida.

Como panorama geral, observamos que as quatro grandes agências ligadas à pesquisa e à pós-graduação – CPNq, CAPES, Finep e FAPESP – têm sistemas razoavelmente institucionalizados e com um número de previsibilidade nos seus resultados. Em contraste, agências setoriais ou instituições não especializadas em apoio à ciência têm procedimentos mais erráticos.

Existe uma permanente e ubíqua disputa de poder entre burocracia e academia. Estes conflitos podem ocorrer, em sua forma mais pura, nas decisões reivindicadas simultaneamente por pesquisadores e administradores sem formação científica. Contudo, é mais freqüente a intermediação das instâncias técnicas das agências. Enredam-se neste litígio, pessoas que há muito tempo lidam com assuntos de fomento à ciência e também uma categoria algo indefinida de pesquisadores, que se vinculam à burocracia das agências e têm suas lealdades e identidade profissionais repartidas entre a academia e a agência a que pertencem.

Adiante examinaremos os procedimentos de três das quatro maiores agências, deixando a CAPES para um capítulo à parte, face à existência de um estudo substancial sobre seu processo de avaliação. Como mencionado, os procedimentos destas agências estão, em boa medida, sintonizados com as práticas usuais e respeitadas nos países cientificamente maduros. Algumas, no entanto, estão aquém do que seria desejável e mesmo possível no nosso nível de desenvolvimento científico. De fato, propostas para seu aperfeiçoamento são freqüentes e muito ruído é produzido na sua discussão.

11. Os Comitês Assessores do CNPq

Tanto por ser a mais antiga agência de fomento à ciência quanto por seu papel oficial de coordenador do Sistema Nacional de Ciência e Tecnologia, o CNPq tem uma posição muito central.

Até meados dos anos 70, seus métodos iniciais de triagem de projetos não diferiam dos adotados usualmente: baseavam-se em juízos informais de um grupo limitado de cientistas transformados em administradores, que se valiam, quando necessário, de pareceres de colegas. Como resultado de uma reformulação interna, foram criados os Comitês Assessores (CAs), compostos de grupos permanentes formados por cientistas com mandatos periódicos e correspondendo a cada área do conhecimento. Desde então, o funcionamento do CNPq tem sido definido em termos da divisão de responsabilidades com os membros dos CAs. Além disto, existe ainda um Conselho Científico e Técnico que, em princípio, traça a política geral do Conselho e seleciona os integrantes dos comitês assessores.

Os CAs têm total poder decisório na avaliação das propostas. Sua palavra é praticamente soberana. Dentro dos orçamentos disponíveis, o CAs selecionam as melhores propostas e deixam suas indicações para serem implementadas pela máquina do CNPq.

O funcionamento desse sistema merece várias observações. Em primeiro lugar, os CAs não se envolvem muito com os detalhes financeiros dos projetos, que ficam por conta da administração do CNPq. Isso, em princípio, parece correto, embora uma melhor articulação entre os CAs e a administração pudesse ser útil em muitos casos.

A total discreção dos CAs na alocação dos recursos tende a tornar os seus pareceres excessivamente lacônicos e pouco elucidativos. O CA não se sente obrigado a justificar-se perante o CNPq. A indicação dos seus membros é o endosso da confiança que *ex ante* desfrutam.

De fato, a autonomia dos CAs deriva de um princípio saudável, que certamente deve ser defendido. Todavia, na prática, essa autonomia tem um aspecto negativo: uma certa desatenção para com os candidatos acerca das justificativas tomadas. Não esqueçamos o caráter educativo dessas comunicações. Decisões tomadas a portas fechadas e cujas razões não são mais adiante explicitadas têm menor legitimidade.

Como qualquer outro sistema que manipula recursos substanciais, o CNPq é objeto de críticas. Não cabe aqui proceder a uma avaliação do CNPq, mas tão-somente identificar pontos controvertidos de seu sistema de avaliação.

Muitas queixas referem-se à formação de igrejinhas e ao elitismo. Ainda que o sistema de rotação bianual dos consultores reduza a formação de grupos de interesse e de sistemas fechados, o fato de o próprio sistema gerar novos nomes para as substituições pode ter os seus vícios.

As panelinhas e igrejinhas são acusações difíceis de aceitar ou refutar. Em uma comunidade científica de porte limitado como a nossa, realmente não há muitos figurantes destacados. É inevitável que os mesmos nomes reapareçam constantemente. Não se pode *a priori* afirmar a existência de bons nomes excluídos ou de incompetentes participando das benesses da instituição. Sabemos que em maior ou menor grau isso ocorrerá; só não sabemos com que freqüência.

Mas afirma-se que competência é função de oportunidades. Cria-se um círculo vicioso ao excluir-se os que não se apresentam tão bem nos concursos. Por não serem beneficiados com recursos, não adquirirão a experiência necessária para competir em igualdade de condições.

As maiores dúvidas quanto ao funcionamento do CNPq, todavia, não estão no funcionamento interno dos CAs, mas nos pontos de tangência destes comitês com o restante da instituição. Não restam dúvidas de que os CAs têm autoridade; talvez devessem mesmo estar mais atentos às justificativas de suas decisões. A dúvida se refere a uma excessiva compartimentalização das suas esferas de influência: poder total sobre a quota de recur-

sos que recebem e nenhuma influência sobre a política científica que gera as regras do jogo e determina as quotas e rubricas onde se pode gastar. Em um sistema amadurecido e bem azeitado isso não seria problema maior. Contudo, em um estágio de transição e redefinições como o que prevalece hoje, o número de arestas e pontos duvidosos é excessivo. Mais ainda, essa divisão estanque é artificial e disfuncional, atingindo detalhes como divisão de gastos entre bolsas e auxílios etc. Quem lida com as propostas está em posição privilegiada para discutir políticas de apoio à sua própria área.

Um ponto muito delicado nesta convivência é à medida que os CAs são utilizados para o financiamento da ciência brasileira. Claramente, o CNPq não canaliza todos os seus recursos desta forma, operando ainda outras categorias em que as decisões são geradas pela sua própria administração. Todavia, essa distribuição entre programas internos e CAs é inevitavelmente arbitrária e conseqüentemente, controvertida. Os funcionários da instituição disputam com os assessores da comunidade a administração dos recursos do órgão.

Os comitês assessores do CNPq produzem periodicamente as *Avaliações e Perspectivas*. Trata-se de uma coleção de avaliações, cobrindo todas as áreas do conhecimento. Não são avaliações de programas ou grupos individuais, mas do estágio de desenvolvimento da área como um todo, com eventuais comentários que individualizam alguns exemplos. Observa-se, de área para área, uma inevitável flutuação na qualidade desses ensaios.

12. A tradição da FAPESP

O sistema da Fundação de Amparo a Pesquisa de São Paulo tem uma boa reputação no que tange à sua estrutura e funcionamento. Inicialmente, as propostas são enviadas para consultores externos (da comunidade científica) para que apresentem por escrito seus comentários. Acompanhadas de seus respectivos pareceres, estas propostas são então apreciadas, em seu conjunto, por um comitê de consultores da área, convocado pela FAPESP. Deste grupo emana a decisão final de aprovar ou recusar um projeto.

Note-se a lógica da divisão de trabalho entre os "pareceristas" e o comitê. Na verdade, tratam-se de duas etapas diferenciadas e que merecem atenção e talentos diferentes.

A revisão técnica de uma proposta é mais bem feita quando realizada por cientistas que desenvolvem linha de trabalho suficientemente próxima do tópico da proposta. Cada parecer, idealmente, requer um volume alentado de trabalho, não sendo possível que poucos se encarreguem de rever em profundidade um número grande de propostas. Os pareceres, portanto, não são imediatamente comparáveis entre si, já que são confeccionados por pessoas diferentes. Não há fórmulas fáceis de se comparar juízes.

Por outro lado, um comitê de pares com o nível de eminência acadêmica desejável e com a visão de conjunto requerida, dificilmente poderia dispor de tempo para examinar as propostas uma a uma e entrar nos seus méritos técnicos ou metodológicos. Mas ainda, cada "sub-sub-área" tem os seus meandros e segredos, sendo impossível para qualquer indivíduo — ou mesmo para um pequeno grupo — conhecer a intimidade de cada tema, de modo a identificar méritos ou fraquezas. Assim, inevitavelmente haverá um resíduo considerável de erros e equívocos nesses julgamentos por comitês auto-suficientes em seu trabalho.

O sistema FAPESP, tal como outros existentes alhures, combina o parecer técnico individual com a visão de conjunto do grupo de pares. Inicialmente, os consultores convocados dão seu parecer mais cuidadoso e bem fundamentado sobre cada projeto individualmente. O comitê se reúne mais adiante e louva-se sobretudo nesses pareceres para o julgamento comparativo, tarefa em que maior visão de conjunto é exigida.

13. A avaliação dos técnicos da FINEP

Pela magnitude dos recursos mobilizados, pela total dependência da *big science* brasileira aos financiamentos por ela concedidos e pelo caráter institucional do seu apoio, sobretudo a grupos novos, a FINEP foi provavelmente a agência de maior expressão em C& T na década de 70. Sob seu apoio financeiro estabeleceram-se muitos dos grupos de melhor reputação hoje.

Ao contrário dos sistemas que delegam grande parte do processo de avaliação a pares da comunidade científica, a FINEP utiliza-se predominantemente do seu próprio pessoal técnico. Para isso dispõe de um amplo quadro de pessoas, geralmente recrutadas na comunidade acadêmica.

O procedimento usual consiste em alocar um ou mais técnicos para examinar cada projeto. Esses técnicos preparam a sua análise que é então submetida à diretoria para o chamado enquadramento, a primeira seleção. Em seguida, os projetos enquadrados são objeto de uma análise mais aprofundada, sendo então novamente submetidos à diretoria para aprovação. Em muitos casos, para auxiliar o julgamento interno, dirimir dúvidas ou lidar com casos mais difíceis, são buscados pareceres externos na própria comunidade acadêmica.

A observação não sistemática sugere que a FINEP tende a escolher consultores de excelente reputação profissional. Todavia, não há critérios explícitos para essa escolha, que depende principalmente de decisões das chefias intermediárias. Tratando-se de uma escolha *ad hoc*, não há garantias de que não seja influenciada pelo interesse do técnico de ver confirmadas suas preferências.

Os procedimentos de avaliação de projetos da FINEP foram objeto de uma tese de doutorado (M. Carakushausky, 1982). Nesta tese a autora

toma 71 projetos julgados pela instituição e reexamina o processo de julgamento. Tais projetos – dos quais 45 foram aprovados – representam uma amostra estratificada das diferentes áreas de conhecimento financiadas pela agência.

Como variável dependente, tomou-se a relação entre os recursos pedidos e os concedidos. Assim, obtêm escores maiores aquelas instituições em que os pedidos são atendidos em maior grau.

Com base na literatura da área, 35 critérios de avaliação foram selecionados. Em análise estatística subseqüente, estes foram agrupados em quatro "macrofatores": relevância do projeto, credibilidade da equipe, volume de produção científica e condições científicas e institucionais do projeto.

Sobre a avaliação dos projetos foram entrevistados 17 pesquisadores da Universidade Federal do Rio de Janeiro e 20 analistas da FINEP, a partir do que, foi possível atribuir notas a cada projeto dentro das escalas e dimensões definidas.

Com esses elementos foram realizados exercícios de correlação parcial, associando o nível de financiamento aos escores nos quatro fatores. Em outras palavras, estar-se-ia implicitamente perguntando em que medida os escores gerados através da análise dos projetos e com o auxílio das entrevistas explicariam o nível de financiamento obtido. Se os processos de julgamento da FINEP fossem arbitrários ou excessivamente eivados de ruídos, seria de se esperar uma associação fraca. Pelo contrário, correlações altas significam consistência de julgamento e validação dos critérios extraídos da literatura.

Os resultados confirmam a segunda hipótese. Os coeficientes de correlação para a produção científica dos doutores do grupo e para a relevância do projeto são respectivamente 0,37 e 0,26. (Note-se que o modelo é diferente e portanto, os resultados não são compatíveis com os da NSF, citados anteriormente).

Em outras palavras, verifica-se que obtêm mais recursos, aqueles grupos que exibem currículos com maior número de publicações e que escolhem temas considerados mais relevantes.

Entrando mais nos detalhes da análise, verifica-se que os fatores que explicam a aprovação são diferentes dos que explicam a rejeição. Bons orçamentos e resultados preliminares, por exemplo, ajudam a aprovar, enquanto que revisões bibliográficas pobres e desarticulação na programação de pesquisa do grupo condenam outros projetos.

Esse é um resultado de algum interesse. Não são poucas as acusações que pesam sobre a FINEP quanto à sua arbitrariedade e falta de critérios. Tal estudo pelo menos demonstra que os erros não estão na orientação geral e tampouco poderiam atingir um número preponderante de projetos; todavia, não permite dizer se a margem de erro é aceitavelmente baixa. De fato, não são resultados tão explícitos e de interpretação tão literal quanto os mencionados para a National Science Foundation.

14. Conclusões: avaliação da ciência ou loteria pseudocientífica?

Nas seções anteriores, examinamos um significativo número de sistemas de avaliação de pesquisas dentro e fora do país. Alguns denominadores comuns podem ser extraídos dessa experiência.

Esses sistemas foram pensados para identificar talento, competência e o interesse do trabalho proposto, refletindo as manifestações tangíveis e arrematadas dessas qualidades. Eles não estão, entretanto, voltados para identificar promessas ou pistas que levem a filões de futuro. Dessa forma, não devem ser entendidos como mecanismos de diagnóstico precoce ou de incentivo a grupos imaturos, ainda que com potencial. Para tais casos existem outros mecanismos que abordaremos mais adiante.

Mas, entendidos como mecanismos para comparar desempenho – e não o seu potencial futuro – são instrumentos razoavelmente confiáveis e previsíveis. Há ampla evidência, inclusive há dois exemplos brasileiros, de que a avaliação da pesquisa produz resultados consistentes e fidedignos. Em outras palavras, propostas julgadas por pessoas diferentes – mas tendo em comun uma boa reputação como cientistas – tendem a receber laudos equivalentes ou parecidos.

Pesquisadores bem formados e que contribuem regularmente para a Literatura da sua área tendem a convergir em sua opinião sobre propostas de pesquisa. A percepção dessa intangível e fugidia "qualidade" é compartilhada por aqueles que são capazes de produzir coisas que também exibem essa qualidade. Na verdade, há uma inevitável circularidade em tudo isso. Mas o objetivo desses processos não é identificar os valores eternos da ciência e sim as pesquisas ou propostas que melhor prossigam o trabalho dos cientistas que aí estão.

Verifica-se também que os critérios proclamados como decisivos nos julgamentos, realmente emergem como centrais nos processos de escolha. Isto é, há consistência entre os valores ditos relevantes e aqueles verdadeiramente utilizados.

Todavia, esses julgamentos estão longe de serem perfeitos. Sua margem de erro ou inconsistência é significativa e deve ser objeto de preocupação e cuidados. Na verdade, essa faixa de equívocos e enganos e, até certo ponto, função dos cuidados e precauções que se possa tomar. Os sistemas existentes evoluíram justamente em função desse empenho em reduzir o erro. De fato, parece seguro afirmar que essa margem de erro é pequena o bastante para garantir-lhes ampla superioridade em relação aos sistemas informais, casuais e desestruturados que os precederam.

Mas é forçoso reconhecer que os atuais sistemas continuam aceitan-

do algumas propostas mais fracas do que outras que foram excluídas, isto é, que eles ainda contêm erros. Por outro lado, as decisões envolvem dimensões de julgamento distintas e que merecerão pesos variados dos diferentes juízes. Não há como dirimir um conflito de preferências entre um projeto melhor resolvido, porém menos original e um outro mais criativo, mas com incorreções e asperezas.

Essa busca de aperfeiçoamento dos sistemas de avaliação tem apresentado progressos consideráveis, obtidos por conta de melhores formatos organizativos e soluções mais modernas ou mais sofisticadas. Todavia, não é só isso. Um sistema mais tosco ou rudimentar, porém bem operado, pode ser mais confiável que outro mais sofisticado mas de funcionamento precário. Em particular, não podemos tomar o que foi visto nas seções anteriores sobre os formatos organizativos e concluir sobre os seus méritos relativos a partir daí. No fundo, estamos falando de formas diferentes de acesso a uma caixa preta que é o julgamento individual. A excelência desse julgamento é o que realmente conta, mais do que a sua correta administração.

Notas

1 A matéria que segue apóia-se na série de trabalhos de John Irving e Ben Martin sobre avaliação de pesquisa. Ver Irving e Martin, 1980, 1983a, 1983b, 1983c e Schwartzman e outros, 1982.

2 É ilustrativo registrar as dificuldades da Finep em relação aos seus grandes projetos, os "finepões", que podem determinar a vida ou a morte de departamentos universitários inteiros. Não é por acaso que o processo decisório nesse órgão extravasa os gabinetes e comissões, entrando mesmo em um cenário político e de *lobbies* externos.

3 Em seus últimos trabalhos, Cartter reverte um pouco essa tendência, fornecendo aos juízes informações sobre os cursos.

III. AS AVALIAÇÕES DA CAPES *

Cláudio de Moura Castro e Gláucio A.D. Soares

A publicação, no *Estado de São Paulo*, em meados de 1982, da lista dos 56 cursos de pós-graduação que obtiveram o pior conceito (nota "E") na avaliação da CAPES provocou um momento de grande excitação na comunidade acadêmica brasileira, culminando num debate que ganhou considerável espaço e cores vivas na grande imprensa.

A divulgação oficial da lista dos cursos de melhor conceito na avaliação da CAPES já havia iniciado uma polêmica que extravasou os circunspectos muros da academia. Em seguida, a revista *Playboy*, em artigo sobre a qualidade do ensino, reiteradamente referiu-se a essas avaliações. Mas foram as entrevistas publicadas com os indignados e vociferantes responsáveis pelos cursos considerados como "sem condições mínimas de funcionamento", que levaram a controvérsia a seu clímax. Alguns chegaram a pensar que o sistema de avaliação, como mecanismo de sinalização para o apoio financeiro à pós-graduação, estaria seriamente ameaçado.

Não obstante, tal teste perante o grande público serviu para consolidar a legitimidade do sistema junto às agências financiadoras de ciência e tecnologia e perante a comunidade acadêmica mais sólida. Na verdade, confirmou-se a estabilidade e robustez de um ambicioso sistema de avaliação da pós-graduação brasileira. É de se notar que tais incidentes coincidiram com a fase final de implementação, na CAPES, de uma série de mecanismos vinculando os auxílios a uma boa nota na avaliação. Mais ainda, foram criados dispositivos que condicionavam o auxílio aos bons cursos a compromissos da universidade de definir uma política realista para os seus cursos mais fracos.

---

* Os autores agradecem as críticas e sugestões de João Batista Araújo e Oliveira, Simon Schwartzman, Raimundo Tadeu Corrêa, Lúcia dos Guaranys e Norma Rancid. Cumpre, todavia, eximi-los da responsabilidade pelos possíveis erros e omissões deste trabalho.

gera coeficientes de 0,98, 0,96 e 0,92. As outras correlações da matriz são igualmente elevadas.

[16] É interessante verificar a predominância, na produção da PUC/RJ, das áreas duras, em contraste com o peso das áreas sociais na PUC/SP.

[17] Ver, neste volume, o Capítulo III – Terceira Parte.

[18] Dados do *Boletim Estatístico* MEC/SESu (novembro 1982).

[19] Note-se que o ISI é muito restritivo, só incluindo as ciências "duras"; as áreas sociais e as engenharias, por exemplo, ficam de fora.

[20] Entre 1980 e 1982 o número de artigos nacionais cresceu de 4.810 para 5.947, isto é, em 24%.

REFERÊNCIAS BIBLIOGRÁFICAS

ADLER, 1985 – *The Quest for Technological Autonomy: Computer and Nuclear Energy Policies in Argentina and Brazil.* Berkeley, University of California Press.

ALBUQUERQUE, LYNALDO C., 1985 – *Política de Gestão de Ciência e Tecnologia.* Brasília, SEPLAN/CNPq.

ALLEN, T.J. E OUTROS, 1983 – "Transferring technology to the small manufacturing firm: a study of technology transfer in three countries". Research Policy (Elsevier), 12, 199-211.

ALTBACH, P., 1985 – "The role and murturing of journals in the third world". Scholarly Publishing, Abril, 211-222.

ALVES, JOSÉ UBYRAJARA, 1981 – "Intervenção nos debates subseqüentes à palestra de Zigman Brener: a instituição e os órgãos financeiros, In: S. Schwartzman e outros, 1981, p. 40.

ALVES, RUBENS, 1984 – *Conversas com quem gosta de ensinar.* São Paulo: Cortez Editora.

ANDREWS, FRANK (ed), 1979 – *Scientific Productivity.* Unesco e Cambridge University Press.

ASTIN, A. and L. SOLMON, 1979 – "Measuring Academic Quality: an Interim Report". Change, 11(6), 48-51.

BECHER, TONY, 1983 – *Research Policies and their impact on research.* Trabalho apresentado à "International Conference on Studies of Higher Education and Research Organization". Rosenon, Dallaro, Suécia, 28, mimeo.

BEN-DAVID, JOSEPH, 1974(1971) – *O Papel do Cientista na Sociedade.* São Paulo: Pioneira.

BEN-DAVID, JOSEPH, 1977 – *Centers of Learning: Britain, France, Germany, United States.* McGraw Hill, The Carnegie Comission on Higher Education.

BERNAL, J.D., 1971(1954) – *Science in History.* Cambridge: The M.I.T. Press, 4 volumes.

BLICKENSTAFF, J. e M. MORAVCSIK, 1982 – "Scientific output in the Third World". Scientometrics, v. 4,2.

BOUDON, RAYMOND, 1977 – "The French university system since 1968". Comparative Politics 10, 1, October, 89-119.

BRENER, ZIGMAN, "A instituição e os órgãos financiadores". In: S. Schwartzman e outros, 1981, 35-41.

BRICKMAN, RONALD, 1977 – "French science policy and the changing role of university". Research Policy 6, 2 (Abril), 128-151.

BRICKMAN, RONALD, 1979 – "Comparative approaches to R&D policy coordination". Policy Sciences 11, 1 (agosto), 73-92.

BRICKMAN, RONALD, 1981 – "The comparative political analysys of science and technology". Comparative Politics 13, 4 (Julho), 479-496.

CAPES, 1978 – *Atividades,* Brasília, Ministério da Educação, Coordenação de Aperfeiçoamento do Pessoal de Ensino Superior.

CAPES, 1982 – *II Plano Nacional de Pós-Graduação.* CAPES informa, 4,2.

CARAKUSHAMSKY, M.S., 1982 – *Avaliação de Projetos Científicos e Tecnológicos.* Rio de Janeiro, COPPE, tese de doutorado.

CARTTER. 1966 – *An Assessment of Quality in Graduate Education.* Washington: American Council of Education.
CARVALHO, GUIDO I., 1975 – *Ensino Superior: Legislação e Jurisprudência.* São Paulo: Editora Revista dos Tribunais, vol. III.
CASTRO, CLÁUDIO DE MOURA, s/d – *As Reitorias de Pesquisa e Pós-Graduação: Papéis e Limitações.* Brasília: Capes, mimeo, 8 p.
CASTRO, CLÁUDIO DE MOURA, 1980 – *A Pesquisa no MEC.* Brasília, mimeo.
CASTRO. CLÁUDIO DE M. e GRACELLI, a sair – *O Desenvolvimento da Pós-Graduação.*
CHAGAS, CARLOS, 1973 – "Pesquisa e Universidade". In: Ciência, Tecnologia e Desenvolvimento. Brasília, Câmara de Deputados, Centro de Documentação e Informação, 71-113.
CLARK, BURTON R., 1980 – *Academic Culture.* New Haven, Higher Education Researc Group, documento de trabalho n. 42, 30 pp.
CLARK, BURTON R., 1983 – *The Higher Education System.* Berkeley: University of California Press.
COLE E COLE, 1967 – "Scientific Output and Recognition". American Sociological Review, v. 62.
COLE, COLE E SIMON, 1981 – "Chance and Consensus in Peer Review". Science, v. 214, novembro.
COLLINS, RANDALL, 1979 – *The Credential Society.* New York: Academic Press.
CROSSLAND, M. (editor), 1976 – *The Emergence of Science in Western Europe.* New York: Science history Publications.
DEBRÉ, M., 1976 – "La langue française et la science universelle". La Recherche, n. 72, nov.
DONGHI, TÚLIO HALPERIN, 1962 – *Historia de la Universidad de Buenos Aires.* Buenos Aires: Eudeba.
DUPREE, A. HUNTER, 1957 –. *Science in the Federal Government.* Cambridge, Mass.: Harvard University Press.
ELIAS, NORBERT, 1982 – "Academic Establishments". In: N. Elias, H. Martins e R. Whitley (eds)., 1982.
ELIAS, N., H. MARTINS E R. WITHLEY (eds), 1982 – *Scientific Establishments and Hierarquies.* Dordrecht: D. Reidl Publ. Co.
FUENZALIDA, E., 1971 – *Investigacion Cientifica y Estratificacion Internacional.* Santiago, Ed. Andres Bello.
GARFIELD, E., 1972 – "Citation analysis as a tool in journal evaluation". Science n. 178.
GARFIELD, E., 1983 – "Mapping Science in the Third World". Science and Public Policy, June.
GILPIN, ROBERT, 1968 – *France in the Age of the Scientific State.* New Jersey: Princeton University Press.
GRASBERG, A., 1959 – "Merit rating and productivity in an industrial research laboratory: a case study". RE Transactions on Engineering Management, 1, 31-37.
HALSEY, A. H., 1961 – "The Changing Functions of Universities", in Halsey, Floud, Anderson, Education, Economy and Society. New York: Free Press.
HERZOG, A., 1983 – "Career patterns of scientists in peripheral communities". Research Policy, n. 12.
HIRSCHMAN, ALBERT O., 1970 – *Exit, Voice and Loyalty – Responses to Decline in Firms, Organizations and States.* Cambridge, Mass.: Harvard University Press.
HUTT, P.B., 1983 – "University-Corporate Research Agreement", Techonology and Society (Pergamon Press), vol. 5, 107-118.
ICFU (International Council on the Future of the University), 1982 – "The academic ethics". Minerva, 1-2 (Spring-Summer), 148.

IRVING, JOHN E BEN MARTIN, 1980 – "Internal Criteria for Scientific Choice: an evaluation of research in high energy physics electron aceleratores"; e "The economic effects of big science: the case of radio astronomy". Proceedings of the International Colloquium on Economic Effects of Space and Other Advanced Technologies. Strasbourg, Abril.
IRVING, JOHN E BEN MARTIN, 1983 – "Assessing basic research: the case of the Isaac Newton telescope". Social Studies of Sciences, v. 13, 49-86.
IRVING, JOHN E BEN MARTIN, 1983a – *Research Evaluation in British Science – a SPRU Review.* Universidade de Sussex, SPRU, abril (mimeo).
IRVING, JOHN E BEN MARTIN, 1983b – *Basic Research in the East and West – a Comparison of Scientific Performance of High-Energy Physics Accelerators.* Universidade de Sussex, SPRU, maio (mimeo).
IRVING, JOHN E BEN MARTIN, 1983c – "Assessing basic research: some partial indicators of scientific progress in radio astronomy". Research Policy, 12, 2.
JOHNSON, R., 1979 – Program to Promote Interaction Between Government, Universities and Industry in the U.K. Seminário sobre "Interação entre governo, pesquisa acadêmica e industrial e atividades de desenvolvimento". Wollogong, Austrália.
KUHN, THOMAS S., 1962 – *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.
LANCASTER, F.W. E M.B. CARVALHO, 1982 – "O cientista brasileiro publica no exterior: em que países, em que revistas, sobre que assuntos". Ciência e Cultura, n. 34(5).
LAWRENCE, J.D. E K.C. GREENE, 1980 – *A question of quality: the higher education ratings game.* HERI Report n.5. Washington, American Association for Higher Education.
LEDERBERG, J., 1983 – "Industrial funding for university research?" Industrial Research Development, sept.
LEDGER, MARSHALL, 1983/4 – "Business and Universities: changing relations". Economic Impact, 92, 37-41.
LYNTON, ERNEST A., 1983 – "As universidades de hoje: uma crise de objetivos". Revista Brasileira de Estudos Pedagógicos, 77, 64.
MAC'ROBERTS, M. E MAC'ROBERTS, B., 1982 – "A reevaluation of Lotka's Law of scientific productivity". Social Studies of Science, v. 12.
MARIN, F., 1976 – *Evaluative Bibliometrics.* Springfield, ed. Computer Horizons.
MASON, STEPHEN S., 1956 – *A History of the Sciences.* New York: MacMillan Publishing Co.
MEC (Ministério da Educação e Cultura), 1975 – *Plano Nacional de Educação* (2ª edição). Brasília, Ministério da Educação e Cultura.
MENDES, CÂNDIDO E CLÁUDIO DE MOURA CASTRO (eds), 1984 – Qualidade, Expansão e Financiamento do Ensino Superior Privado. Rio, ABM/EDUCAM.
MERTON, ROBERT K, 1957 – *Social Theory and Social Structure.* New York: The Free Press.
MERTON, ROBERT K., 1970(1938) – *Science, Technology and Society in Seventeenth Century England.* New York: Harper & Row.
MINTZBERG, HENRY, 1979 – *The Structuring of Organizations.* Englewood Cliffs: Prentice Hall.
MOREL, R. E C. MOREL, 1977 – "Um estudo sobre a produção científica brasileira segundo os dados do ISI". C. Informação, 6(2).
NARIN, F, 1976 – *Evaluative Bibliometrics: the use of publications and citation analysis in the evaluation of scientific activity.* New Jersey: Computer Horizon.
NATIONAL SCIENCE FOUNDATION, 1981 – *Academic Science, 1972-1981.* NSF 81-326, Wasington, D.C.
NATIONAL SCIENCE FOUNDATION, 1983 – *University-Industry Research Rela-*

tionships. Washington, D.C.: NSF.
NETHERLANDS ORGANIZATION FOR SCIENTIFIC RESEARCH ZWO, 1980 – *University Research.* The Hage: Ministry of Education and Science.
NEEDHAM, JOSEPH, 1969 – *The Grand Tritation: Science and Society in East and West.* Londres: Allen & Urwin.
NOBLE, DAVID, 1977 – *America by Design.* New York: Knopf.
NUNES, MARCIA B. M., N. V. X. SOUZA E S. SCHWARTZMAN, 1982 – "Pós-graduação em engenharia: a experiência da COPPE". In: Schwartzman (ed), 1982.
OLIVEIRA, JOÃO B. A., 1984 – *Ilhas de Competência: Carreiras Científicas no Brasil.* São Paulo, Brasiliense.
OLIVEIRA, JOÃO B. A. E S. SCHWARTZMAN, 1980 – "Autonomia Universitária". Estudos e Debates, 3.
OLIVEIRA, JOÃO B. E ROGER WALKER. "Tecnologias no ensino e na administração universitária". In: Cândido Mendes e Cláudio de Moura Castro (eds), 1984, 86-116.
ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT, 1972-1974 – *The Research System,* vols. 1-3. Paris, OECD.
PARSONS, TALCOTT, 1951 – "The Institutionalization of Scientific Investigation", in Bernard Barber e Walter Hirsch, The Sociology of Science. New York: Free Press.
PAULINYI, ERNO, 1984a – "Dispêndios nacionais de ciência e tecnologia". Revista Brasileira de Tecnologia, 15, 2 (março-abril).
PAULINYI, ERNO, 1984b – "Os dispêndios nacionais em C&T e o produto interno bruto". Revista Brasileira de Tecnologia, set./out.
PIRSIG, ROBERT, 1975 – *Zen and the Art of Motorcycle Maintenance.* New York, Bantan Books.
PLUCIENNIK, MOYSES A, 1981 – *Papéis organizacionais em uma instituição de pesquisa.* Publicação interna n. 106. São José dos Campos, Instituto Tecnologico da Aeronáutica, outubro.
PRESIDÊNCIA DA REPÚBLICA, SECRETARIA DE PLANEJAMENTO, 1981, 1982, 1983, 1984 – *Orçamento da União para Ciência e Tecnologia.* Brasília, CNPq/Seplan.
PRESIDÊNCIA DA REPÚBLICA, SECRETARIA DE PLANEJAMENTO, 1982a – *Setor Produtivo Estatal – Dispêndios em Ciência e Tecnologia,* 1979/82. Brasília, CNPq/Seplan.
PRESIDÊNCIA DA REPÚBLICA, SECRETARIA DE PLANEJAMENTO, 1984a – *Recursos do Tesouro Estadual – 1983 – Orçamento de Ciência e Tecnologia, Estados e Territórios.* Brasília, CNPq/Seplan.
PRICE, D. SOLA, 1963 – *Little Science, Big Science.* New York: Columbia University Press.
PRICE, D. SOLA, 1972 – *Science and Technology: distinctions and relationships".* In: B. Barnes, (ed), Sociology of Science. Harmondsworth: Penguin Books.
RATTNER, H. (coord), C. Machline e C.O. Bertero, 1983 – *Política e Administração de Tecnologia.* São Paulo, Blucher Editora.
REIS, FABIO W.; H. P. GODINHO E I. M. CAMPOS, 1981 – *Uma Avaliação Preliminar da Pós-Graduação e Pesquisa na UFMG – Indicadores, Perfis e Problemas Especiais.* Belo Horizonte: Universidade Federal de Minas Gerais.
ROCHE, MARCEL E Y. FREITES, 1982 – "Produccion y flujo de informacion cientifica en un pais periferico americano (Venezuela)". Interciência, vol. 7, set.out.
ROKKAN, STEIN, 1974 – "Entries, voices, exits: towards a possible generalization of the Hirschman model". Social Sciences Information, 13, 1, 39-53.
ROSE, K. AND C. ANDERSEN, 1970 – *A Rating of Graduate Programs.* Washington: American Council of Education, 1970.

ROY, RUSTUM, 1983 – "Peer review of proposals – rationale, practice and performance". Bulletin of Science and Technology, Pergamon Press.
SAVIANI, DEMERVAL, 1984 – *Ensino Público e Algumas Falas Sobre Universidade.* São Paulo, Cortez Editora.
SCHWARTZ, M., K. PAVITT E R.ROTHWELL, 1982 – *The Assessment of Governmental Support for Industrial Research: Lessons from a Study of Norway.* Universidade de Sussex, SPRU, outubro (mimeo).
SCHWARTZMAN, S. E OUTROS, 1979 – *Formação da Comunidade Científica no Brasil.* Rio e São Paulo: FINEP/Cia. Editora Nacional.
SCHWARTZMAN, S. E OUTROS, 1981 – *Administração da Atividade Científica.* Brasília, FINEP/CNPq.
SCHWARTZMAN, S. (ed), 1982 – *Universidades e Instituições Científicas no Rio de Janeiro.* Brasília, CNPq.
SCHWARTZMAN, S., 1983 – *The Quest for University Research.* Rio de Janeiro: Iuperj, Série Estudos, n. 19.
SCHWARTZMAN, S., 1984a – *Coming full circle – for a reappraisal of University Research.* Rio de Janeiro: Iuperj, Série Estudos, n. 31.
SCHWARTZMAN, S., 1984b – "A política brasileira de publicações científicas e técnicas". Revista Brasileira de Tecnologia, maio-junho.
SCHWARTZMAN, S., 1985 – *Organização e Desempenho da Pesquisa Científica no Brasil (relatório final da parte brasileira do Projeto ICSOPRU).* Rio de Janeiro, IUPERJ, manuscrito.
SIPEGEL-ROSING, INA E DEREK DE SOLLA PRICE (editores), 1977 – Science, *Technology and Society.* Los Angeles: Sage Publications.
SMITH, GRAHAM, 1983 – *Carta a Social Studies of Science*, vol. 13.
SOLMON, L., 1977 – "The Catter Report on the leading schools of education, law and business". Change, 1977, 48(4), 44-48.
STANKIEWITZ, E., 1979 – *Social Process of Utilization of Scientific Knowledege – a theoretical essay.* Oslo, Institute for Studies in Research and Higher Education, mimeo.
UFMG (Universidade Federal de Minas Gerais, 1982 – *Pós-Graduação e Pesquisa na UFMG, 1981: Uma avaliação preliminar.* Belo Horizonte, UFMG, Conselho de Pós-Graduação.
VELHO, LEA E J. KRIEGE, 1984 – "Publication and Citation practices of Brazilian agricultural scientists". Social Studies of Science, v. 14, 4562.
VESSURI, H., 1982 – "Las relaciones entre Universidad y Aparato Produtivo". Acta Científica Venezolana (Caracas), 33, 4-14.
WEINBERG, ALVIN M., 1961 – "Impact of large-scale science on the United States. Science, 21, July, 1614.
WEINSTEIN, GILKA, 1981 – "Intervenção no debate posterior à palestra de Zigman Brener". In: S. Schwartzman e outros, 1981, p. 40.
WEICK, KARL, 1979 – *The Social Psychology of Organizing.* Reading, Mass.: Addison-Wesley Publ. Co.
WEISS, CAROL H., 1980 – "Knwoledge creep and decision accretion". Knowledge 1, p. 392.
WILSON, LOGAN, 1979 – *American Academics: Then and Now.* New York: Oxford University Press.
ZUCKERMAN, H., 1967 – "Nobel laureates in science". American Sociological Review, n. 32.